10th Postal Economics Conference on E-commerce, Digital Economy and Delivery Services
Institut D'Economie Industrielle Université Toulouse, March 29-30, 2018

Estimation of demand for unaddressed mail in the UK: entry and exit customer analysis using firm level data*

Frederique Feve[1], Jean Pierre Florens[1], Leticia Veruete-McKay[2] and Soterios Soteri[2]

1. Introduction

Postal mail volumes in the UK, similar to other advanced countries, exhibit quite different trends amongst different types of traffic. For example, the advance of electronic communications and e-commerce have led to a structural decline in letter volumes and increasing parcel volumes. However, even within these segments of traffic different elements exhibit quite diverse behaviour (see PwC (2013)). In order to make informed business decisions and allocate resources more efficiently it is therefore important that postal operators obtain a better understanding of the key drivers of the demand for different types of mail and, in particular, the extent to which the prices they offer impact customer demand.

Marketing mail in the UK accounts for around one-quarter of total letter revenue and competes with a wide range of alternative media for advertising budgets. However, it is unclear to what extent postal operators should respond to such competition via price or other factors. This paper examines a rich customer data set for unaddressed advertising mail to examine the extent to which changes in price impact the demand for such mail.

In particular, we estimate a number of dynamic demand for mail models for unaddressed advertising mail that takes into account the panel structure of customers' volume and number of contracts at geographic zonal areas since 2007. These models are estimated using parametric and non-parametric techniques and we provide estimates for short and long run price elasticities. We are not aware that these types of demand models have been estimated in the context of postal services for unaddressed mail.

This paper is structured as follows. Section 2 provides an overview of the data and the framework of the dynamic demand models for unaddressed mail. Section 3 contains empirical results based on parametric and non-parametric techniques and section 4 contains a summary and conclusions.

2. Data and dynamic demand models for unaddressed advertising mail

Unaddressed advertising mail is delivered in the UK via its Door to Door (D2D) product range which mainly consists of marketing material (e.g. advertising flyers, brochures and other material contained

---

* The views expressed in this paper are those of the authors and do not necessarily reflect those of their affiliated organizations.

[1] Affiliated to IDEI, University of Toulouse
[2] Royal Mail Group

in an envelope that does not include a recipients name and is delivered to all households within a specific geographical area). A key feature of unaddressed mail is that it is relatively quick and easy to distribute to all households and business delivery points in the UK. Customers (business, government and other) who use unaddressed mail tend to send physical mail communications for promotional and other purposes that target specific postcode or geographic areas.

Our data set consists of weekly customer transactions and we have information on volumes and prices charged by zonal areas and various firm level characteristics; (e.g. whether they are a new customer or have considered advanced booking). From this information we constructed a dataset with the unit of observation being a customer contract for the period July 2007 to May 2013.

Contracts can be booked in advance and the duration of the contract could be up to thirteen weeks. The price charged relates to the month in which the contract was booked with its corresponding volume (number of items sent) in zonal areas chosen. But this follows a non-linear pricing rule, where a discount applies to each contract, and the discount is a function of the total volume in a given contract. For instance, if customers sent up to 25,000 items the price charged is the same across the zonal areas, but if it is above 25,000 item the price varies across zonal areas. But also the higher the volumes of over 25,000 items the lower the prices. The average price was constructed for each month and zonal areas and weighted by the corresponding number of items sent in a given contract. Note that new zonal areas have been introduced in 2011 and 2012.

A panel was constructed where the variables are observed by month t and by pricing zone j (t=1,…, T, j=1,….J), taking into account also changes in the number of pricing zones during the period studied.

We estimated a dynamic demand model based on the following specification:

$$\text{Ln } Y_{tj} = \varphi_j (\ln Y_{t-1\,j},\ \ln P_{tj}, \ln X_{tj},\ D_{tj}) + \varepsilon_j + u_{tj} \qquad (1)$$

where

- Y_{tj} represents the level of demand of unaddressed mail UK in month t and pricing zone j. This could be measured by either the number of contracts or the number of items sent in the month
- P_{tj} is the average price
- X_{tj} refers to macroeconomic variables (various variables were considered such as UK retail sales and production)
- D_{tj} refers to dummy variables (to help account for the impacts of temporary and permanent shocks to changes in the definition of pricing zones in November 2011 and April 2012)
- ε_j denotes heterogeneity zonal areas term
- $u_{t,j}$ denotes an error term

Different models were considered depending on the choice of φ_j taking into account the following options:

a) Endogeneity of P_{tj} or not,
b) Heterogeneity on zonal areas j or not,

c) Different macroeconomic variables in X_{tj} and
d) Temporary or permanent shocks in pricing zonal areas in .

The common feature of the model specification (1) is the autoregressive structure which creates a dynamic specification in the different models estimated and allows us to estimate both short and long term price elasticities.

3. Estimated short and long term elasticities

The price elasticity can be computed by the following expression, taking into account the price dependence on the number of items sent:

$$\frac{\partial \varphi_j}{\partial \ln P_{tj}}$$

This is the short term price elasticity but depends on the explanatory variables and can be time varying in our sample.

While the long term price elasticity is computed by:

$$\frac{\frac{\partial \varphi_j}{\partial \ln P_{tj}}}{1-\frac{\partial \varphi_j}{\partial \ln Y_{t-1\,j}}}$$

This expression computes the value of the long run effect of a price variation if all the variables remain identical to their values at the observation (t j).

Different dynamic demand models were estimated, first using parametric techniques in section 3.1 and second using non-parametric techniques in section 3.2.

3.1. Parametric linear models

Considering expression (1), several linear models were estimated using parametric techniques. First we assumed that φ_j is linear and not dependent on j as follows:

$$\ln Y_{tj} = \alpha + \beta \ln Y_{t-1\,j} + \gamma \ln P_{tj} + \delta' \ln X_{tj} + \lambda' D_{tj} + u_{tj} \qquad (2)$$

another variant of the model was an interaction between zonal area, dummy and prices:

$$\ln Y_{tj} = \alpha + \beta \ln Y_{t-1\,j} + \sum_{j=1}^{J} \gamma_l \left(\ln P_{tl}\ x\, 1|\,(j=l)\right) + \delta' \ln X_{tj} + \lambda' D_{tj} + u_{tj} \qquad (3)$$

These models were estimated by Ordinary Least Squares (OLS) (assuming prices are exogenous) and by the instrumental variable method using Two Stage Least Squares (2SLS) (allowing for endogeneity in prices).

In both models, we have the dynamic parameter β and a short term price elasticity γ identified for all zonal areas together as in (2) and γ_l, specific for each zonal area j as specified in (3). Long run price elasticities are then equal to $\frac{\gamma}{1-\beta}$ and $\frac{\gamma_l}{1-\beta}$ respectively.

Linear model by pricing zone

We adopted the same specification as (2), but this time estimated models where all parameters depend on the zonal area j, as follows:

$$\ln Y_{tj} = \alpha_j + \beta_j \ln Y_{t-1\,j} + \gamma_j \ln P_{tj} + \delta_j' \ \ln X_{tj} + \lambda_j' D_{tj} + u_{tj} \qquad (4)$$

This equation was estimated for each individual zonal area using OLS and 2SLS. Computations of the price elasticity are the same as before except that all the β, δ and λ may differ by zonal area. The heterogeneity between zonal areas is captured by the different parameters. This model is very general but each zonal area model is estimated using a smaller sample of observations. The variance of the demand is also different for each zonal area.

Panel structure

In addition, the dynamic demand model was estimated with a panel structure (e.g. with fixed effects), allowing to capture the heterogeneity in the pricing zonal areas across customer contracts in the term, ε_j, as in (1).

3.2 Non- parametric estimation of demand for unaddressed mail

We estimated local linear models using non-parametric estimation techniques. Note that these models did not introduce the heterogeneity term nor the endogeneity in price, since this would require more complex methods of estimation which are at the frontier of economics research (see Fève and Florens, 2014) and lie outside the scope of this particular paper.

We then consider the following type of models:

$$ln\, Y_{tj} = \alpha(\xi_{tj}) + \beta\,(\xi_{tj}) ln\, Y_{\,t-1\,j} + \gamma(\xi_{tj})\ ln\, P_{tj} + \delta(\xi_{tj})'\ ln\, X_{tj} + \lambda(\xi_{tj})'\, D_{tj} + u_{tj} \qquad (5)$$

This model differs from the linear models by the dependence of the parameters to some chosen variables ξ_{ij}. We restrict the dimension of ξ_{ij} in our application and consider two case. In the first, ξ_{ij} depends only on P_{tj} and in the second case ξ_{ij} depends on both Y_{t-1j} and P_{tj}. In the latter case the short term price elasticity (ξ_{tj}) is therefore a function of both the price and the lagged explained variable.

The estimation method is an extension to kernel methods and is called a "local polynomial methods". Basically, for a given value of ξ we minimize:

$$\sum_{tj} (\ln Y_{tj} - \alpha(\xi) - \beta(\xi) \ln Y_{t-1\,j} - \gamma(\xi) \ln P_{tj} - \delta(\xi)' X_{tj} - \lambda(\xi)' D_{tj})^2 \ \ K\left(\frac{\xi - \xi_{tj}}{h}\right) \qquad (6)$$

where K is a Naradaya Watson kernel (in our example, the density of a gamma distribution[1]) up to a scaling factor h.

This computation is replicated for all the observed values of ξ which generate the results.

[1] $K(u) = e^{\frac{-u^2}{2}}$

The model (6) was estimated in a nonparametric way by pooling all the data similarly as in formula (2).

4. Empirical results

Using statistical criteria, we narrowed down our results from all the various models we estimated using different estimation methods (parametric and non-parametric techniques); different macro economy variables (GDP, volume of retail sales and the index of output services) and aggregation levels (pooling all zonal areas together and by individual zonal area).

In general, we found that models containing macro economy variables improved the explanatory power of the models and the statistical validity of the estimated price elasticities. In contrast, the IV models yielded estimated price elasticities that had low t-statistics and exhibited substantial volatility when the instrument list (W) was varied and when this was extended to include more variables and lags of these variables the results started to converge onto the OLS estimates. It was also the case that the results using volume data yielded less robust results than those using the number of contracts. This is due to the volume data exhibiting a high degree of volatility and are more difficult to explain. This may be due to data collection issues or may simply reflect high volatility in customer demand.

A summary of our empirical results for the estimated unaddressed advertising price elasticities are reported in table 1 and further details on the estimated models yielding these results are contained in the appendix.

The estimated unaddressed advertising price elasticities reported in table 1 show that the long-term elasticities lie in the range -0.6 to -1.5 and are around twice the magnitude of the short-term estimated elasticities. This property is due to the autoregressive parameter being positive and smaller than unity in all the estimated models (see appendix) and is consistent with the economic interpretation of price effects cumulating across time and stabilizing in a stationary model.

The results also show that the estimated elasticities differ for pricing zones and that zones A and B tend to be higher in absolute terms.

Table 1. Estimated price elasticities for UK unaddressed mail

Dependent variable	*Linear models* Estimated using parametric techniques		*Non-linear model* Estimated using non parametric techniques	
Number of contracts	***Pooling all zonal areas***		***Pooling all zonal areas***	
Zonal area prices	Short term	Long term	Short term	Long term
A	-0.38 to -0.71	-0.83 to -1.46	-0.56	-0.92
B	-0.33 to -0.65	-0.78 to -1.35	-0.43	-0.76
C	-0.31 to -0.63	-0.67 to -1.29	-0.79	-1.86
D	-0.29 to -0.60	-0.63 to -1.25	n.s.	n.s.
E	-0.29 to -0.60	-0.62 to -1.24	n.s.	n.s.
Volumes sent	***Pooling all zonal areas together***			
Zonal area prices				
A	-0.69 to -0.92	-0.97 to -1.31	Models not estimated	
B	-0.57 to -0.80	-0.81 to -1.14		
C	-0.53 to -0.75	-0.75 to -1.08		
D	-0.49 to -0.71	-0.68 to -1.01		
E	-0.47 to -0.69	-0.66 to -0.98		
Number of contracts	**Pooling individual zonal areas**			
Zonal area prices				
A	-0.60 to -0.64	-0.94 to -1.03	Models not estimated	
B to E	n.s.	n.s.		

Notes:* Denotes statistically significant at 5% significance level. ** Denotes statistically significant at 10% significance level. n.s. denotes results not statistically significant. Figures in parenthesis are standard errors for price coefficients. Estimation period covers June 2007 to May 2013.

5. Conclusions

The range of estimated price elasticities for unaddressed advertising mail is quite wide (-0.6 to -1.5) with a mid-value, in absolute terms of around unity. This suggests that the extent to which price can be used generate higher financial returns or respond to competitive pressures is not completely clear and we would recommend that further analysis be undertaken to narrow the range of estimates. In particular, two areas that could be explored further are the choice of macro economy variables in the econometric models and the pattern and nature of the dynamic impacts by, for example, adopting richer dynamic models using both parametric and non-parametric estimation techniques.

References

Fève, F. and J.P. Florens (2014), 'Nonparametric analysis of Panel Data Models with Endogenous Variables', *Journal of Econometrics*, 181(2):151-164.

Hendry, D. (1995), 'Dynamic Econometrics, Advanced Texts in Econometrics. Oxford University Press.

PwC (2013). The outlook for UK mail volumes to 2023. http://www.royalmailgroup.com/sites/default/files/The%20outlook%20for%20UK%20mail%20volumes%20to%202023.pdf

Wooldridge, J., 2010, *Econometric analysis of cross section and panel data*. MIT press.

Appendix

Table 2. Estimated elasticities for UK unaddressed mail

Dependent variable: Number of contracts									
	Linear models: pooling together price zonal areas				*Liner models: pooling by individual price zonal areas*				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
Lagged dependent variable (Y_{t-1})	0.54*	0.54*	0.51*	0.52*	0.36*	0.38*	0.65*	0.52*	0.30*
Lagged dependent variable (Y_{t-2})	×	×	×	×	×	×	×	×	0.24**
Zonal area prices									
A	-0.38*	-0.40*	-0.65*	-0.71*	-0.60*	×	×	×	×
B	-0.33*	-0.35*	-0.60*	-0.65*	×	-0.61	×	×	×
C	-0.31*	-0.33*	-0.57*	-0.63*	×	×	-0.94	×	×
D	-0.29*	-0.31*	-0.55*	-0.60*	×	×	×	-0.24	×
E	-0.29*	-0.31*	-0.55*	-0.60*	×	×	×	×	-0.19
Index of production	×	×	×	×	×	×	×	×	×
Retail sales	2.17*	×	1.78	×	2.41	×	1.39	×	1.64
GDP	0.92*	1.23*	×	×	×	×	×	×	×
Index of services	×	×	1.85*	2.15*	×	1.97**	1.41	2.48*	2.78
Estimation period June 2007 to May 2013									
Number of observations	355	355	355	355	71	71	71	71	71
R^2 adjusted	0.80	0.79	0.81	0.81	0.54	0.60	0.63	0.65	0.63

Notes:* Denotes statistically significant at 5% significance level. ** Denotes statistically significant at 10% significance level. Figures in parenthesis are standard errors for price coefficients. × denotes that the variable was not included in the estimation. All explanatory variables include a constant, dummy variables and logarithmic values of the explanatory variables noted in the table above.

Table 3. Estimated elasticities for UK unaddressed mail

Dependent variable: Volumes sent				
	Linear models: pooling together price zonal areas			
	(1)	*(2)*	*(3)*	*(4)*
Lagged dependent variable	0.29*	0.30*	0.38*	0.30*
Zonal area prices				
A	-0.69*	-0.92*	-0.29*	-0.90*
B	-0.57*	-0.80*	-0.19	-0.78*
C	-0.53*	-0.75*	-0.15	-0.73*
D	-0.49*	-0.71*	-0.12	-0.69*
E	-0.47*	-0.69*	-0.10*	-0.67*
Index of production	1.05	1.50*	1.67*	1.19**
Retail sales	5.36*	4.59*	3.15	5.41*
GDP	×	×	×	×
Index of services	×	×	×	×
Estimation period June 2007 to May 2013				
Number of observations	355	355	355	355
R^2 adjusted	0.60	0.61	0.55	0.62

Notes:* Denotes statistically significant at 5% significance level. ** Denotes statistically significant at 10% significance level. Figures in parenthesis are standard errors for price coefficients. × denotes that the variable was not included in the estimation. All explanatory variables include a constant, dummy variables and logarithmic values of the explanatory variables noted in the table above.